JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2007, **87,** 401–404 NUMBER 3 (MAY)

WITHIN-TRIAL CONTRAST: WHEN IS A FAILURE TO REPLICATE NOT A TYPE I ERROR?

THOMAS R. ZENTALL AND REBECCA A. SINGER

UNIVERSITY OF KENTUCKY

Vasconcelos, Urcuioli, and Lionello-DeNolf (2007) report the results of five experiments that fail to replicate the results of our within-trial contrast study (Clement, Feltus, Kaiser, & Zentall, 2000) and suggest that our results may represent a Type I Error. We believe that this conclusion is not warranted because (a) there is considerable evidence in support of the effect and (b) the amount of training that they gave to their pigeons prior to test may not have been sufficient to observe the effect reliably. We suggest that when sufficient training is provided, reliable contrast can be found.

Key words: work ethic, simultaneous discrimination, within-trial contrast, Type I Error, overtraining, key peck, pigeons

Clement, Feltus, Kaiser, and Zentall (2000) trained pigeons on two simultaneous hue discriminations. On half of the trials, a single peck (low effort) was required to obtain one of the simultaneous discriminations. On the remaining trials, 20 pecks (high effort) were required to obtain the other. Following extended training, when the pigeons were given a choice between the two S+ stimuli, they showed a strong preference for the one that in training required greater effort to obtain. This effect was described as a form of contrast involving the greater improvement in value from the high-effort response to the S+ that followed than from the low-effort response to the S+ that followed. A model was proposed in which an S+ stimulus preceded by a relatively more aversive event would be preferred over a similar S+ stimulus preceded by a less aversive event (Zentall, Clement, Friedrich, & DiGian, 2006).

Vasconcelos, Urcuioli, and Lionello-DeNolf (2007) report the results of five experiments that fail to replicate the results of our within-trial contrast study. Their results raise doubts about the reliability of the phenomenon. Vasconcelos et al. propose that the lack of reliability suggests that our finding may be attributable to a Type I Error—the 1 chance in 20 that the difference found was, in fact, not different from chance. Vasconcelos et al. cite Davison and Nevin (2005) who argue that positive findings are more easily published than negative findings.

Address correspondence to: Thomas R Zentall, Department of Psychology, University of Kentucky, Lexington, KY 40506-0044 (E-mail: zentall@uky.edu).

doi: 10.1901/jeab.2007.04-07

Were the Clement et al. (2000) study the only one to report such an effect, the Vasconcelos et al. (2007) conclusion might be warranted, but the fact that there have been a number of other empirical studies that have found a similar contrast effect (Clement & Zentall, 2002; DiGian, Friedrich, & Zentall, 2004; Friedrich, Clement, & Zentall, 2005; Friedrich & Zentall, 2004; Kacelnik & Marsh, 2002; Klein, Bhatt, & Zentall, 2005; Marsh, Schuck-Palm, & Kacelnik, 2004; Pompilio & Kacelnik, 2005; Pompilio, Kacelnik, & Behmer, 2006; Singer, Berry, & Zentall, 2007; Singer & Zentall, 2007) suggests a different conclusion. The studies that have found evidence in support of a similar contrast effect represent 13 independent experiments. If our findings were attributable to a Type I Error, it would suggest that there should have been approximately 250 experiments conducted (but mostly not reported) that failed to find the effect. Although such a record of failed experiments is possible because there is no way to estimate the number of unpublished negative results, it would seem unlikely.

LEVEL OF TRAINING

We, too, have been concerned with the ability of Urcuioli and his collaborators (and others) to observe this within-trial contrast effect, and we have explored several possibilities. One of these is the amount of training given prior to testing. In several of our experiments and in the Vasconcelos et al. (2007) experiments, pigeons were given 20 overtraining sessions (following acquisition of the two simultaneous hue discriminations). A

problem in designing these contrast experiments is that it is difficult to know how much training is needed prior to testing. Although there is a clear measure of acquisition of the two simultaneous discriminations, there is no independent measure of the acquisition of the (Pavlovian) relation between the prior (relatively aversive) event and the S+ that follows. In our initial experiments (Clement et al., 2000; Clement & Zentall, 2002; DiGian et al., 2004; Friedrich et al., 2005), we provided the pigeons with an arbitrary number of sessions of overtraining (20) following acquisition of the two simultaneous discriminations. At the time, we thought that 20 sessions would surely be sufficient to establish a strong relation between the prior event and the simultaneous discrimination that followed. However, in later research, we began to monitor the development of the contrast effect by inserting probe trials during training (Friedrich & Zentall, 2004) and found that the contrast effect was quite slow to develop. For example, Singer et al. (2007) found that a reliable preference for the S+ stimulus that followed the less-preferred prior event did not emerge until the pigeons had had 30 sessions of overtraining (following attainment of criterion on the simultaneous discriminations). Similarly, Friedrich and Zentall (2004) found that a reliable preference did not emerge until the pigeons had experienced 60 sessions of overtraining. More recently, in an experiment similar to that of Singer et al., a reliable preference emerged only after more than 40 overtraining sessions (Singer & Zentall, 2007). Thus, extensive training appears to be needed to reliably observe this contrast effect. In fact, we too have had some difficulty in replicating this contrast effect with only 20 overtraining sessions (Klein & Zentall, 2002). Thus, the association between the prior event and the simultaneous discrimination appears to develop rather slowly and that association is critical for the development of the reported contrast effect.

OTHER FAILURES TO FIND THE WITHIN-TRIAL CONTRAST EFFECT

There have been other attempts to obtain the contrast effect that we have reported. One of these studies involved rats running up inclined ramps (high effort) versus level ramps (low effort, Armus, 1999). Given the willingness of rats to work for access to a running wheel in the absence of any other reinforcer, we suspect that the inclined ramp does not provide a sufficiently aversive experience for contrast effects to be found.

Two other studies that have failed to find a contrast effect involved rats pressing levers in which the effort required to press the lever was manipulated (Armus, 2001; Jellison, 2003). In neither study were the rats overtrained. Furthermore, in both studies, different flavored pellets followed low and high effort and the rats were later tested in a Y or T maze for their acquired flavor preference (one flavor was placed in each arm and the rats had to learn in which arm each flavor was placed). Thus, any acquired flavor preference during training may have been reduced because these test trials, in effect, were training trials involving a new task, and it is not clear how long after lever-press training the potential flavor preferences would be expected to persist, especially in the new, maze context. In addition, the Y and T mazes may not be particularly sensitive for detecting modest flavor preferences because in such contexts, rats are known to show spontaneous alternation (Dember & Fowler, 1958).

In contrast, Jellison (2003) found that over 70% of the rats (12 of 17) that showed a flavor preference preferred the flavor associated with the high-effort response. Although this effect was not statistically reliable, it was a reasonably large effect in the direction expected for a contrast effect.

Why the Variability in Results?

The question remains why we have been able to observe the contrast effect with less than 30 sessions of training (Clement et al., 2000; Clement & Zentall, 2002; DiGian et al., 2004; Friedrich et al., 2005), whereas Vasconcelos et al. (2007) have not. Our best guess is that the lower limit on the amount of training needed to observe this contrast effect is about 20 sessions and at that level the reliability of the effect is questionable. However, when sufficient training is provided, a reliable contrast effect has been found (see Table 1).

If, in fact, the amount of training provided in those studies is near the threshold required to observe the contrast effect, then other procedural differences may influence whether or not the effect is observed. For example,

Table 1

Studies with Pigeons that Reported and that Failed to Report a Within-Trial Contrast Effect.

Observed Contrast?	Number of Overtraining Sessions			
	20	30	40	60
Yes	Clement et al. (2000) Clement & Zentall (2002) DiGian et al. (2004) Friedrich et al. (2005)	Singer et al. (2007)	Singer & Zentall (2007)	Friedrich & Zentall (2004)
No	Vasconcelos et al. (2007) Klein & Zentall (2002)			

Note – Several studies with rats provided no overtraining and failed to find a contrast effect (Armus, 1999, 2001; Jellison, 2003). However, other studies with insects (Pompilio, Kacelnik, & Behmer, 2006), starlings (Kacelnik & Marsh, 2002; Marsh, Schuck-Paim, & Kacelnik, 2004; Pompilio & Kacelnik, 2005) and humans (Klein, Bhatt, & Zentall, 2005) also provided no overtraining and did report a contrast effect.

a potentially important moderating variable could be individual color preferences. We have found that in a simultaneous hue discrimination, prior to training, pigeons sometimes prefer one hue over another and that preference often shows up on probe test trials in which the two preferences are in competition (Zentall, Dorrance, & Clement, 1999). Were pigeons to have strong hue preferences, it would not necessarily eliminate the development of contrast effects but it would likely increase the between-subject variability of choice and limit the degree to which those contrast effects would appear in preference tests.

CONCLUSIONS

(1) The within-trial contrast effect reported by Clement et al. (2000) has been replicated under various conditions and in more than one laboratory.
(2) A factor that appears to be important in observing the contrast effect is the amount of original training.
(3) If insufficient training is given, the effect may be variable, especially if other factors, such as color preferences, come into play.

REFERENCES

Armus, H. L. (1999). Effects of response effort on secondary reward value. *Psychological Reports, 84*, 323–328.

Armus, H. L. (2001). Effect of response effort on the reward value of distinctively flavored food pellets. *Psychological Reports, 88*, 1031–1034.

Clement, T. S., Feltus, J., Kaiser, D. H., & Zentall, T. R. (2000). ''Work ethic'' in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review, 7*, 100–106.

Clement, T. S., & Zentall, T. R. (2002). Second-order contrast based on the expectation of effort and reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 64–74.

Davison, M., & Nevin, J. A. (2005). On science and the discriminative law of effect. *Journal of the Experimental Analysis of Behavior, 83*, 85–92.

Dember, W. N., & Fowler, H. (1958). Spontaneous alternation behavior. *Psychological Bulletin, 55*, 412–428.

DiGian, K. A., Friedrich, A. M., & Zentall, T. R. (2004). Reinforcers that follow a delay have added value for pigeons. *Psychonomic Bulletin & Review, 11*, 889–895.

Friedrich, A. M., Clement, T. S., & Zentall, T. R. (2005). Reinforcers that follow the absence of reinforcement have added value for pigeons. *Learning & Behavior, 33*, 337–342.

Friedrich, A. M., & Zentall, T. R. (2004). Pigeons shift their preference toward locations of food that take more effort to obtain. *Behavioural Processes, 67*, 405–415.

Jellison, J. L. (2003). Justification of effort in rats: Effects of physical and discriminative effort on reward value. *Psychological Reports, 93*, 1095–1100.

Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour, 63*, 245–250.

Klein, E. D., Bhatt, R. S., & Zentall, T. R. (2005). Contrast and the justification of effort. *Psychonomic Bulletin & Review, 12*, 335–339.

Klein, E. D., & Zentall, T. R. (2002). [Failure to replicate the ''work ethic'' effect in pigeons.] Unpublished raw data.

Marsh, B., Schuck-Palm, C., & Kacelnik, A. (2004). State-dependent learning affects foraging choices in starlings. *Behavioral Ecology, 15*, 396–399.

Pompilio, L., & Kacelnik, A. (2005). State-dependent learning and suboptimal choice: When starlings prefer long over short delays to food. *Animal Behaviour, 70*, 571–578.

Pompilio, L., Kacelnik, A., & Behmer, S. (2006, March 17). State-dependent learned valuation drives choice in an invertebrate. *Science, 311*, 1613–1615.

Singer, R. A., Berry, L. M., & Zentall, T. R. (2007). Preference for a stimulus that follows an aversive event: Contrast or delay reduction? *Journal of the Experimental Analysis of Behavior, 87*, 275–285.

Singer, R. A., & Zentall, T. R. (2007). [Within-trial contrast is found only after extensive training.] Unpublished raw data.

Vasconcelos, M., Urcuioli, P. J., & Lionello-DeNolf, K. M. (2007). Failure to replicate the ''work ethic'' effect in pigeons. *Journal of the Experimental Analysis of Behavior, 87*, 383–399.

Zentall, T. R., Clement, T. S., Friedrich, A. M., & DiGian, K. A. (2006). Stimuli signaling rewards that follow a less preferred event are themselves preferred: Implications for cognitive dissonance. In E. A. Wasserman, & T. R. Zentall (Eds.), *Comparative cognition: Experimental explorations of animal intelligence* (pp. 651–667). New York: Oxford University Press.

Zentall, T. R., Dorrance, B. R., & Clement, T. S. (1999). Differential inhibition and stimulus generalization cannot account for value transfer in simultaneous discrimination learning by pigeons: Reply to Aitken. *Animal Learning & Behavior, 27*, 494–496.

Received: January 5, 2007
Final acceptance: January 13, 2007